Modeling Spectral Cubes in $\mathcal{AIPS}$

Eric W. Greisen

May 21, 2014

Abstract

$\mathcal{AIPS}$ has done Gaussian fitting along the x-axis of image cubes with task `XGAUS` since the 1980s. That task has recently been overhauled to be much easier to use and much more capable. In like fashion, new tasks `ZEMAN` and `RMFIT` have been developed. The former fits the standard leakage and scaling terms for Stokes V cubes, including a new option to do this for each of the Gaussians found by `XGAUS`. The latter fits polarization models to Stokes Q and U cubes, using the output of Faraday Rotation Measure Synthesis ($\mathcal{AIPS}$ task `FARS`) to assist with initial guesses. The models can contain multiple components each with a polarization flux, angle, rotation measure, and rotation measure "thickness." The present memo will describe the functions of these tasks in some details with numerous graphical examples.

1 Introduction

`XGAUS` was written, and re-written, a long time ago in the hopes that converting a spectral cube, with many spectral channels at each spatial location, into one or more sets of images of peak, center, width, and integral would simplify the data presentation. One could even hope that the separate Gaussians fit might have some separate physical reality. The task could fit a single Gaussian to each spatial position without much human guidance, but detailed human interactivity was required for more complex spectra. Use of the task involved an endurance contest between the human and the power company, requiring both luck and skill on the part of the human. The task had no way to back up and fix fits generated at previous spatial positions and either wrote out all of the fit parameters in images or discarded all work done on the cube. Clearly, this was not a desirable situation and would-be users tended to prefer fitting routines in other packages such as `GYPSY`.

There has always been some interest in fitting rotation-measures to polarization images taken at different frequencies. Until the advent of the Jansky VLA however, the images available for fitting rotation measures were at a modest number of scattered frequencies and could generally be fit by rather simple algorithms. This is no longer the case and better fitting algorithms were required. Initially there was considerable interest in a method known as Faraday Rotation Measure Synthesis.[1] The method was implemented in $\mathcal{AIPS}$ by Leonid Kogan as task `FARS` with supplementary tasks and additions by the author. However, it was found that the "synthesized beam width" in this method was too wide to separate most scientifically interesting cases, even if the data were taken to frequencies as low as 1 GHz. Data from lower frequencies would be needed to produce better component discrimination, but most radio sources are completely depolarized at these low frequencies. A more direct fitting method is therefore required.

The author undertook a re-write of `XGAUS` in order to make it a much more useful and user-friendly task and to provide a software framework for a new task to do polarization fitting. The suggestion was then made to create a task to fit standard Zeeman-splitting parameters as well. The following Memo attempts to describe in detail the function of all three of the tasks that have resulted. This memo is based in the `31DEC14` version of these tasks and does not attempt to describe differences with the `31DEC13` version, in which these changes were first made.

---

[1] Brentjens, M. A. and de Bruyn, A. G. 2005, *Astron. & Ap.*, 1217.

There are also tasks in AIPS which modify *uv*-data files to add model components both spatially and spectrally. In addition there are tasks which add polarization and spectral models to images. These will be described in a final section.

2 Gaussian fitting: XGAUS

Gaussian fitting is the process of determining the peak brightness, center position, and width of a number of one-dimensional Gaussians, plus a baseline, which best matches a one-dimensional slice of input data. XGAUS will do this for every voxel in a three-dimensional image cube, taking each row as the one-dimensional slice. It then produces images over the second and third input axes of the fit parameters. Normal usage has the first axis of the cube as a spectral axis, either in frequency or velocity units, and the second and third axes are celestial coordinates. This usage will be assumed in the later discussion, but it is not required. Mathematically, the function fit is

$$T(x, y, z) = A + Bx + \sum_{1}^{\text{NGAUSS}} T_i e^{-\alpha(x-x_i)^2/\sigma_i^2} \qquad (1)$$

where x is the coordinate along the first axis of the image cube, $A(y, z) + B(y, z)x$ is the spectral baseline, $T_i(y, z)$ are the peak values of each of the NGAUSS components, $x_i(y, z)$ are the center channels of the components, $\sigma_i(y, z)$ are the full widths at half maximum of the components and $\alpha = 4 \ln(2)$ to impart this meaning to the σ_i.

2.1 Inputs

The usual INNAME *et al.* adverbs define the cube to be fit and the OUTNAME *et al.* adverbs define the (eventual) output name. The INVERS adverb controls which XG table is used by the task, with ≤ 0 meaning a new table. BLC and TRC define the pixel ranges to be used in the current execution, where BLC(1) and TRC(1) control the spectral channels that will be fit and BLC(2), TRC(2), BLC(3), and TRC(3) control the area in celestial coordinates over which the fitting is done. If you are resuming work on a pre-existing XG table, the current fitting area must equal of be inside of that specified when the table was created. Adverbs YINC and ZINC control the stride taken in the first pass through the cube; a second pass will then fit all voxels not fit in the first pass. FLUX controls with spectra will be fit during this execution; all positions with 3 consecutive channels averaging above FLUX will be fit. The initial guess for the linear baseline will be made with channels BLC(1) to BLC(1) + BCHAN − 1 and channels TRC(1) − ECHAN + 1 through TRC(1). Good baselines are usually fit even if BCHAN and ECHAN are left zero. DOOUTPUT controls what files are written — this may be changed interactively so leave it zero at this point. Set DOTV = 2 to use TV menus to prompt you. Even when fitting only 1 Gaussian component, it is best to watch what is happening so you should never set this adverb false. DORESID controls whether residuals are plotted on the fit spectra; such plots often provide clues when more Gaussians are needed for the best fits. In general it is best to set PIXRANGE to zero to see the full range of image values, but, if you are fitting weak Gaussians in the presence of very strong ones, you might wish to cut off the highest values. (This may make it harder to set initial guesses however since you can't point at the peak of the clipped component.) Set LTYPE to your favorite type of labeling, set PIXVAL to zero to see all positions being fit, and leave NITER zero since 100 is more than enough iterations. Set NGAUSS to the number of Gaussians to be fit in this execution. It may be changed if you re-start on a pre-existing XG table since those tables contain room for the maximum number of Gaussians allowed (8 at this writing). RMSLIMIT is an upper limit for the rms of a fit before the fit is viewed as "failed" which causes the TV and interaction to be turned back on after you have turned it off. You should get a good idea of an appropriate value from your initial uses of XGAUS or from your knowledge of the noise in your data cube.

A recommended strategy, when creating an XG table, would be to set the fitting area (BLC(2), TRC(2), BLC(3), and TRC(3)) to something bigger than the full region with usable signals that you might ever want to fit. The task begins by creating this table and populating each row with the largest average brightness
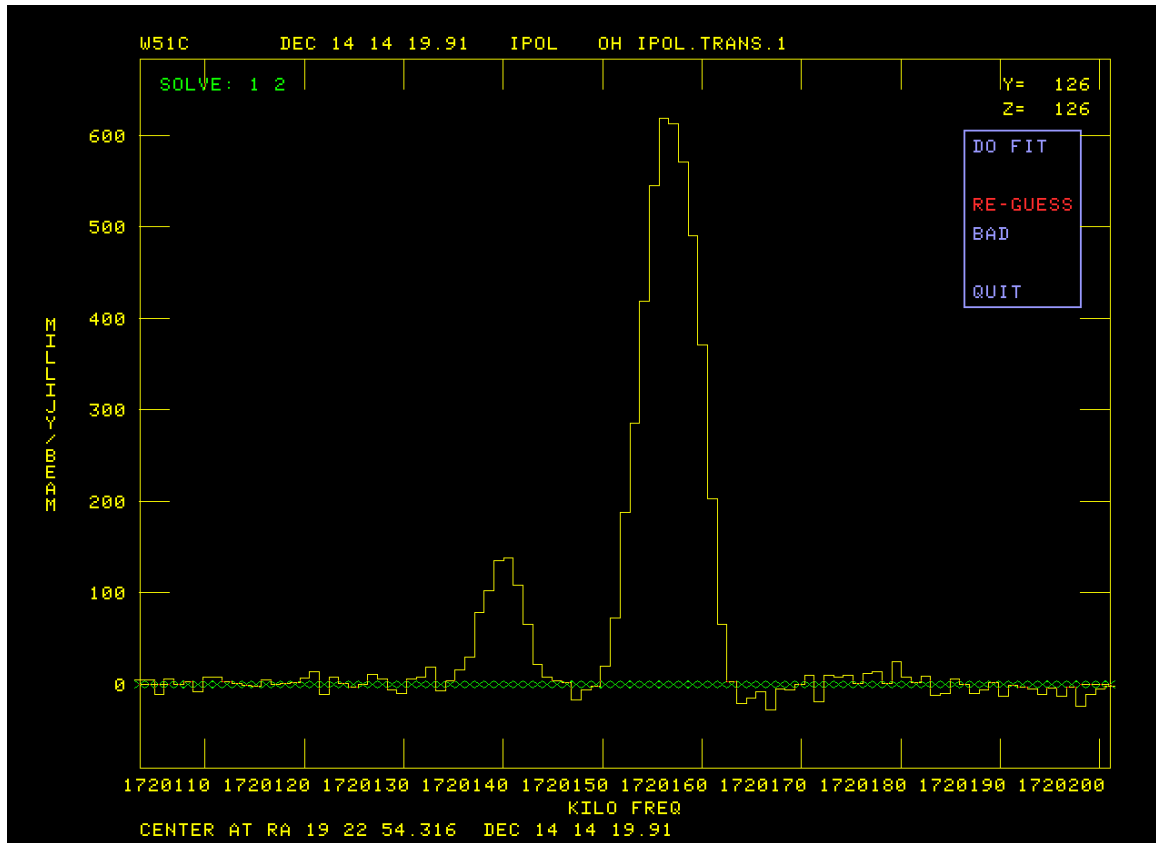
Figure 1: First OH spectrum to fit, initial "guess" is all zero.

over three consecutive channels in the corresponding row. Then it reads the table every `YINC` rows and `ZINC` planes and, for those with an average brightness greater than `FLUX` attempts a fit. Your interaction with this fit will be described below. After the first pass, the task loops over every row and plane fitting those positions which have enough brightness and which have not already been fit. Finally, after all pixels above `FLUX` have been fit, the task goes into an "edit" mode. It constructs images of each fit parameter and the integral flux of the Gaussian and of the uncertainties in these parameters. You may view these images, select positions explicitly or by their parameter values or rms and revisit the fits of those positions. This stage will be described in detail below.

At any time you may exit the task and then re-start it using the same `XG` table. Good reasons for doing this include fitting smaller regions with each pass using the appropriate number of Gaussians for that region. Doing small regions which will have similar parameter values helps a great deal with the initial guessing done by the task (mostly using the previous solution). You might also fit the cube initially with a high value of `FLUX` and then re-start with a lower value to extend the areas fit.

2.2 Fitting

The fitting process starts with a plot of the spectrum with the data and axis labels in graphics channel one (usually yellow) and the initial guess as X's in graphics channel two (usually green). Then you are offered a menu of options, either in your $\mathcal{AIPS}$ terminal window (`DOTV = 1`) or, as shown in the present figures, on the TV (`DOTV = 2`. The first spectrum to be fit (from an OH 1720 MHz maser source) is illustrated in Figure 1 showing that the first guess for more than one Gaussian is not useful. The menu that appears at this point is
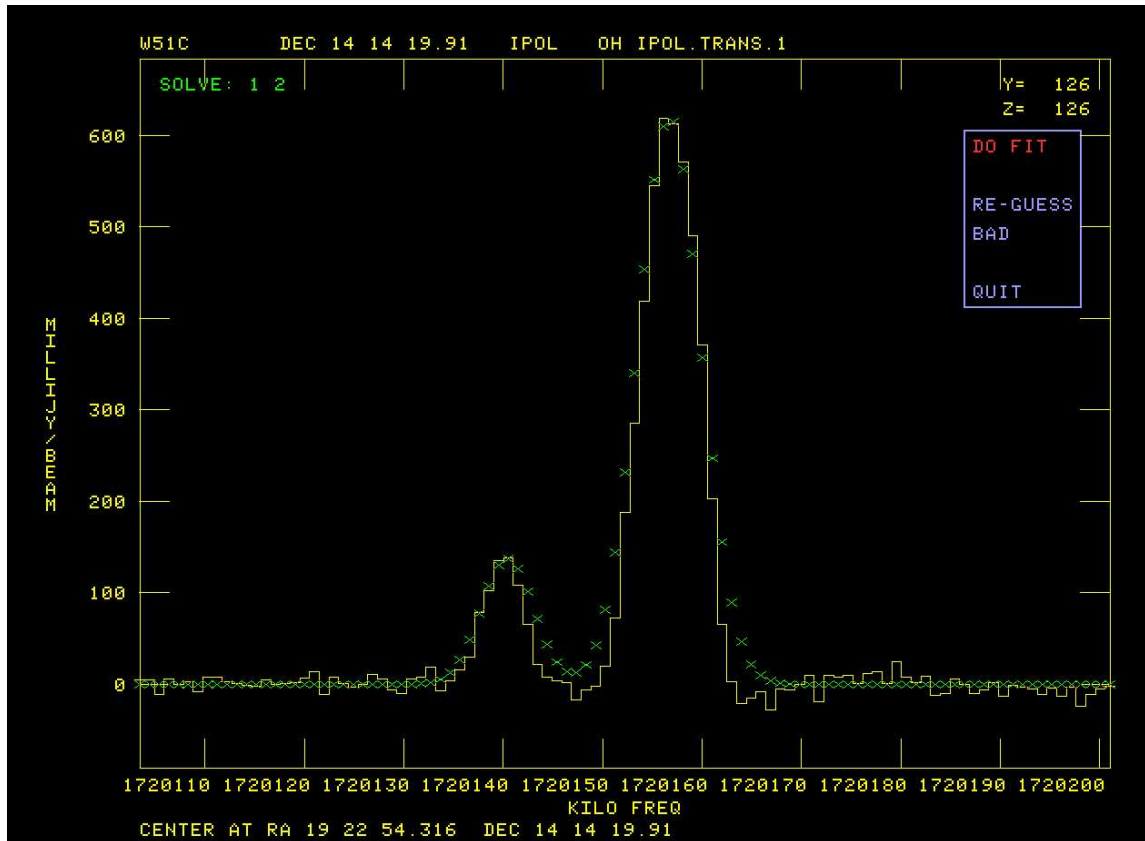
Figure 2: First OH spectrum to fit: better initial "guess" has been entered.

DO FIT	
RE-GUESS	E
BAD	B
QUIT	Q

You select a menu option by moving the cursor to the desired option with the mouse and registering that move with the TV by clicking the left mouse button. The selected menu item will change color as shown in the figure. If you press TV "button" D at this point (actually keyboard character D), helpful information about the selected item will appear on your terminal window. If you press one of TV "buttons" A, B, or C (actually keyboard characters A, B, C), the selected function will be performed. The option to QUIT (or Q on the terminal) causes the task to quit at this point. You may re-start later. The option BAD (B on the terminal) will mark this position as failed and go on to the next position. The option DO FIT will cause the task to attempt the non-linear fit with the current initial guess. The selected option in Figure 1 is RE-GUESS which causes the task to prompt you first to *"Position cursor at center & height of Gaussian component 1"*. Move the cursor to the peak of component 1 and press any TV button. This selects the peak value and center of component 1. Then the task prompts you to *"Position cursor at half-width of Gaussian component 1"*. Move the cursor horizontally to the approximate position of the half-power point of component 1 and press any TV button. The horizontal position of the cursor then sets the initial guess of the full width of the component. These prompts are then repeated for components two through NGAUSS. If you do not want to fit a particular component at this position, move the cursor off the plot before pressing the TV button for that component; it will be omitted.
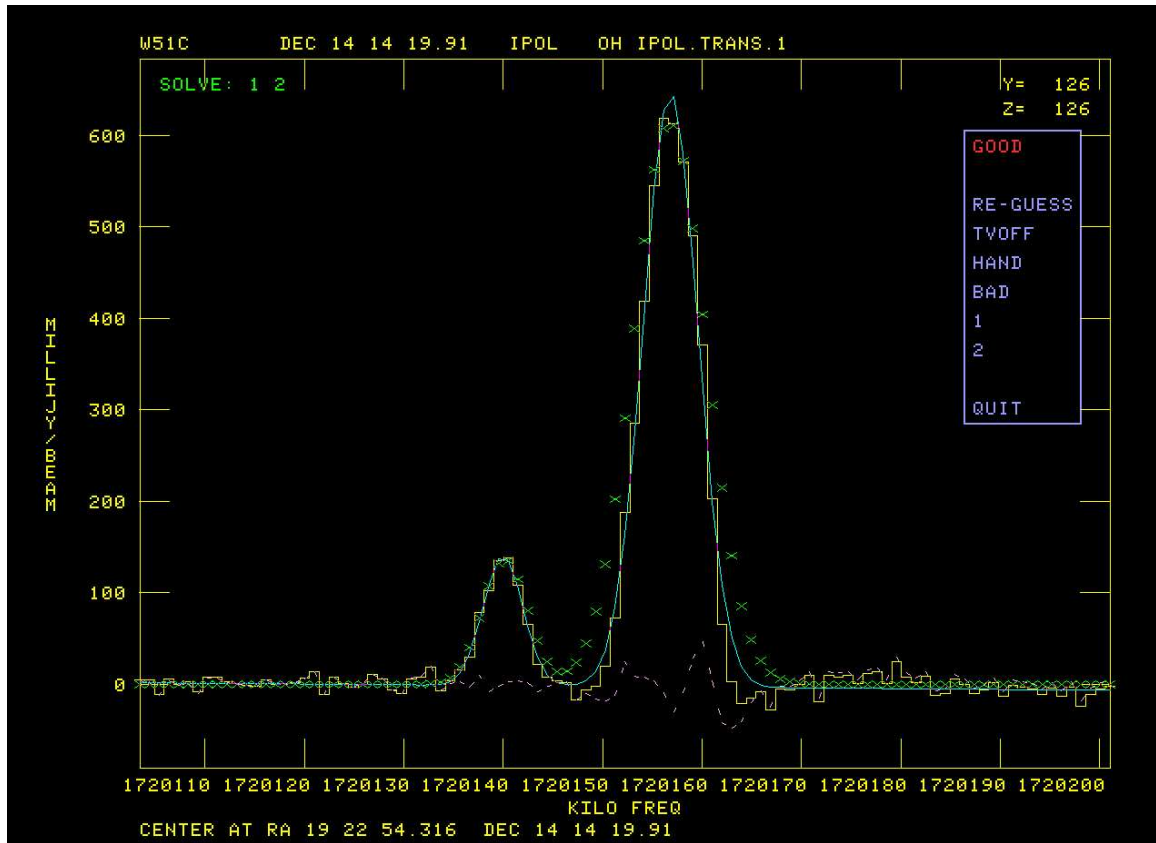
Figure 3: First OH spectrum to fit: good fit obtained.

After the improved initial guess has been entered the plot is changed to show the new guess, as illustrated in Figure 2. This guess is good so the `DO FIT` option is selected. After TV 'buttons" `A`, `B`, or `C` are pressed the task attempts the fit with the current initial guess. Then the plot is changed with the addition of the fit function in graphics plane 4 (usually cyan) and the residual in graphics plane 3 (usually pink). This is illustrated in Figure 3. A different menu appears at this point containing

GOOD	
RE-GUESS	E or R
TVOFF	T
HAND	H
BAD	B
1	1
2	2
QUIT	Q

The option to `QUIT` (or `Q` on the terminal) causes the task to quit at this point. You may re-start later. The option `BAD` (`B` on the terminal) will mark this position as failed and go on to the next position. The option `RE-GUESS` (`E` or `R` on the terminal) will loop back to prompt you for a new guess and repeat the fit. Options `1`, `2`, ..., `NGAUSS` will loop back to plot an initial guess with the selected number of Gaussians. Option `HAND` (`H` on the terminal) will prompt you to enter using the terminal the Gaussian parameters for each component. Enter on one line for each component, the peak value of the Gaussian (in image units), the center (in pixels with respect to the reference pixel), and width (in pixels). `XGAUS` will then repeat the display in Figure 3 to see if you made a good guess. Option `TVOFF` allows you to turn off interactivity, allowing the task to run
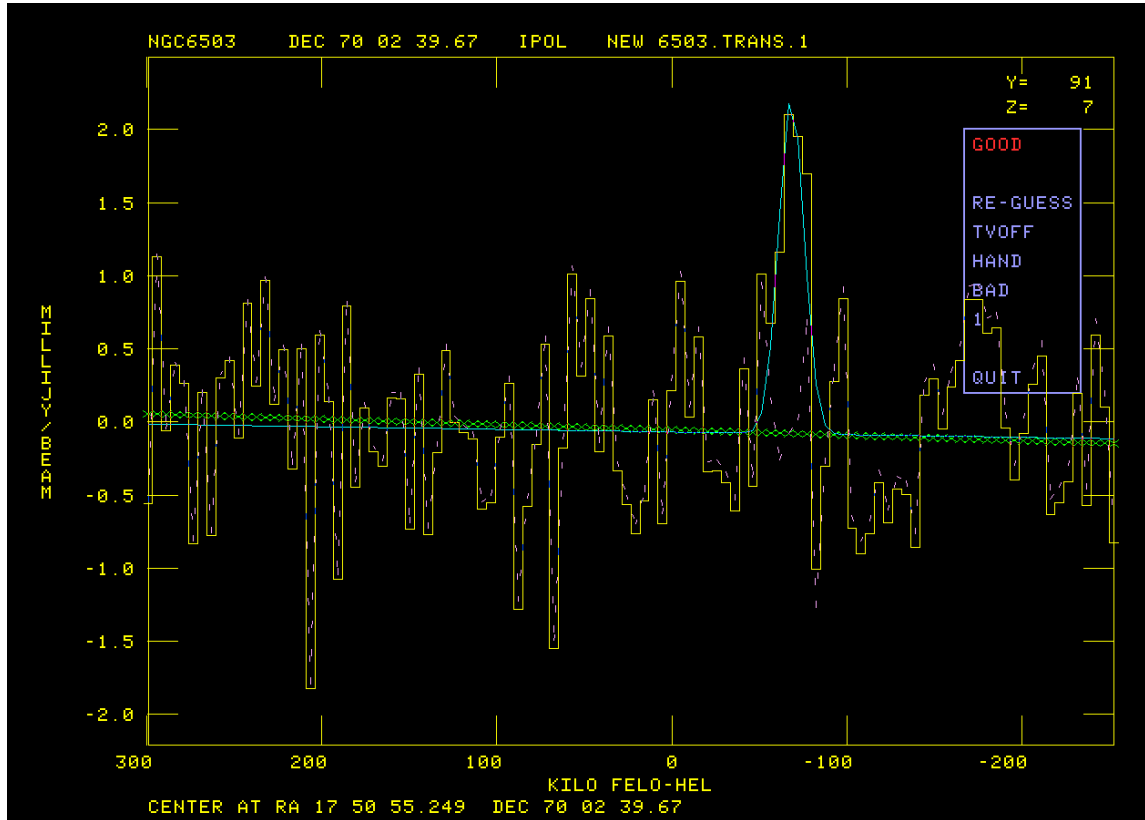
Figure 4: First 1-Gaussian HI spectrum to fit: good fit obtained automatically.

using its own initial guesses until it finds a completely unreasonable solution or one with an rms greater than RMSLIMIT. When that happens, you are shown the offending fit parameters and the task resumes with the plot of Figure 3 to allow you to try to fix things. Option GOOD (and other initial character on the terminal) tells the task that you are (reasonably) happy and that it should go on to the next position.

When fitting only a single Gaussian, XGAUS makes an initial guess based mostly on finding a real peak in the spectrum. This is quite reliable, so turning off the TV interaction may save a great deal of effort, although there will possibly be bad positions to be fixed up in the next stage of this task. An example of an HI cube, on its first position, is shown in Figure 4.

2.3   Editing and output

Eventually all positions selected by FLUX and BLC and TRC will have been fit. At this point, the task computes images of the fit parameters plus the "flux" (area under the Gaussian) and their uncertainties. It then offers a lengthy menu of options which will allow you to view these images and revisit positions that seem to have produced incorrect fits. If NGAUSS > 1, options to swap portions of image n with corresponding portions of image m are also offered. This "edit" menu is illustrated in Figure 5. Note that the size of the OH image is quite small, more interesting images will appear below and also note that, for legibility in all figures in this memo, the $\mathcal{AIPS}$ TV was run with double-sized characters. The titles normally require only one line. The first column of the menu shows the following functions:

There are three kinds of editing implemented here. In the first, the user establishes the parameter extrema which should be viewed as acceptable. The extrema currently set are shown in the title lines. Then, XGAUS may be told to flag all solutions not meeting these criteria, or, more profitably perhaps, to revisit those
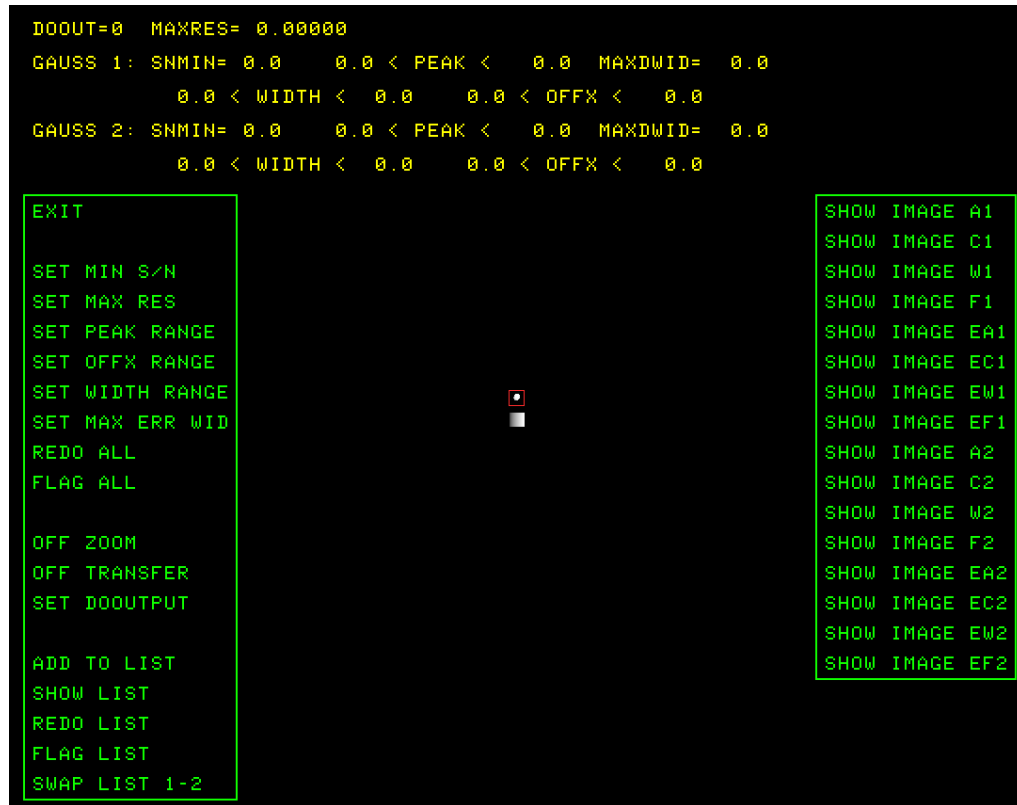
Figure 5: OH maser field, edit view

positions to see if a better fit can be obtained. The other editing methods are similar, but act on a list of pixel positions. These may be entered by typing in values or by clicking on suspect pixel positions in the CURVALUE function described below. The contents of the list may be viewed, the solutions at the positions may be flagged, they may be revisited to attempt for a better solution, or the solutions at the listed positions may be swapped between components n and m. After the flagging, revisiting, or swapping, the list is cleared.

|	
EXIT	Exit XGAUS, writing output images if DOOUTPUT is now > 0.
SET MIN S/N	Set minimum amplitude S/N(s) for okay solutions
SET MAX RES	Set maximum residual for okay solutions
SET PEAK RANGE	Set peak value range(s) for okay solutions
SET OFFX RANGE	Set offset range(s) for okay solutions
SET WIDTH RANGE	Set width range(s) for okay solutions
SET MAX ERR WID	Set maximum error(s) in width for okay solutions
REDO ALL	Re-do all solutions which are not okay
FLAG ALL	Mark bad all solutions which are not okay
OFF ZOOM	Turn of TV zoom
OFF TRANSFER	Turn off black & white and color TV enhancements
SET DOOUTPUT	Increment DOOUTPUT in loop 0-3 — with 1 and 3 causing residual images and 2 and 3 causing parameter images to be written on EXIT
ADD TO LIST	Type in output pixel coordinates to add to list
SHOW LIST	Display coordinates in list
REDO LIST	Re-do solutions for all pixels in list
FLAG LIST	Flag solutions for all pixels in list
SWAP LIST 1-2	Swap solutions for components 1 and 2 for all pixels in list
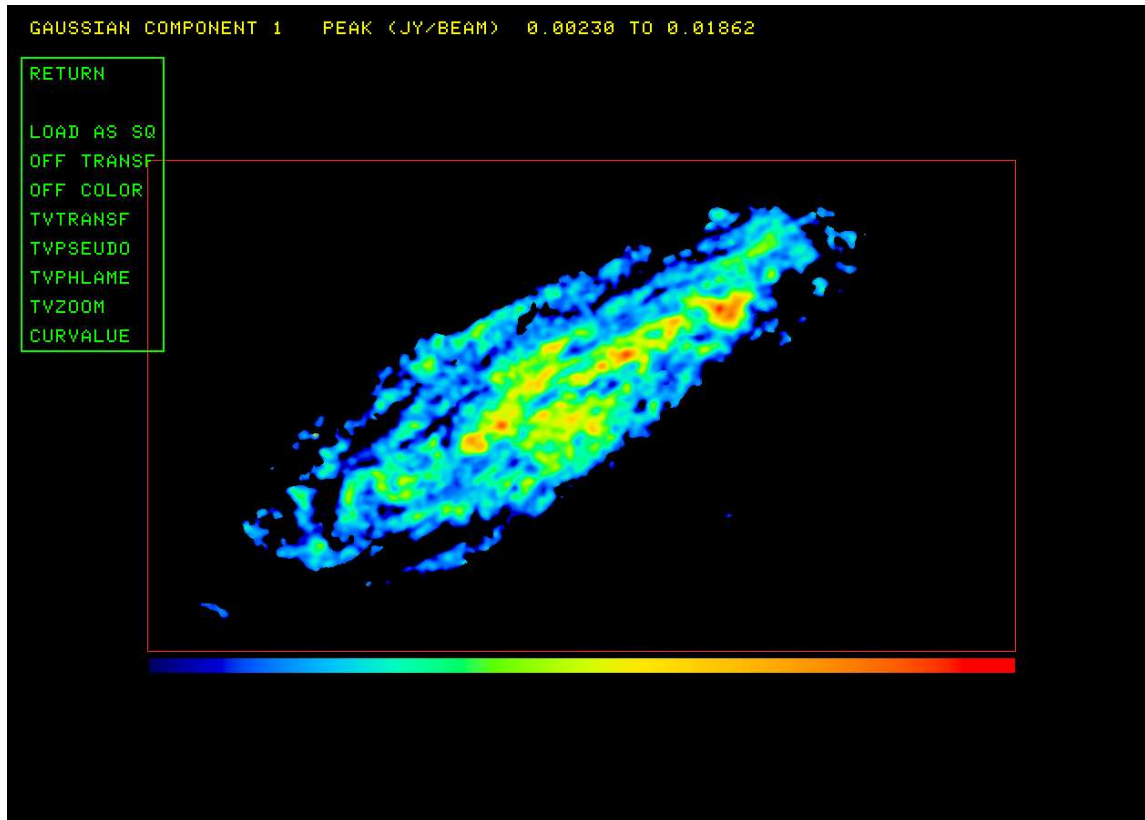
Figure 6: HI galaxy: peak value image interaction

The second (and potentially third and even fourth) menu columns contain `NGAUSS` sets of functions

SHOW IMAGE A1	Enter image interaction with peak value of component 1
SHOW IMAGE C1	Enter image interaction with center pixel of component 1
SHOW IMAGE W1	Enter image interaction with width of component 1
SHOW IMAGE F1	Enter image interaction with "flux" of component 1
SHOW IMAGE EA1	Enter image interaction with uncertainty in peak value of component 1
SHOW IMAGE EC1	Enter image interaction with uncertainty in center pixel of component 1
SHOW IMAGE EW1	Enter image interaction with uncertainty in width of component 1
SHOW IMAGE EF1	Enter image interaction with uncertainty in "flux" of component 1

When you select one of these functions most of the following operations will appear in yet another menu. This menu is illustrated in Figure 6. Only one of the `LOAD AS` options will appear, with the next one in the sequence offered when the current one has been invoked. The `SWAP` n-m options will appear as needed when `NGAUSS` > 1. The `NEXT WINDOW` option appears when needed to display an image too large to fit on the display screen. These options mostly invoke familiar functions from $\mathcal{AIPS}$ to control the `FUNCTYPE` used in loading the image to the display, to enhance the image intensities, to color the enhanced image intensities, and to zoom the display.
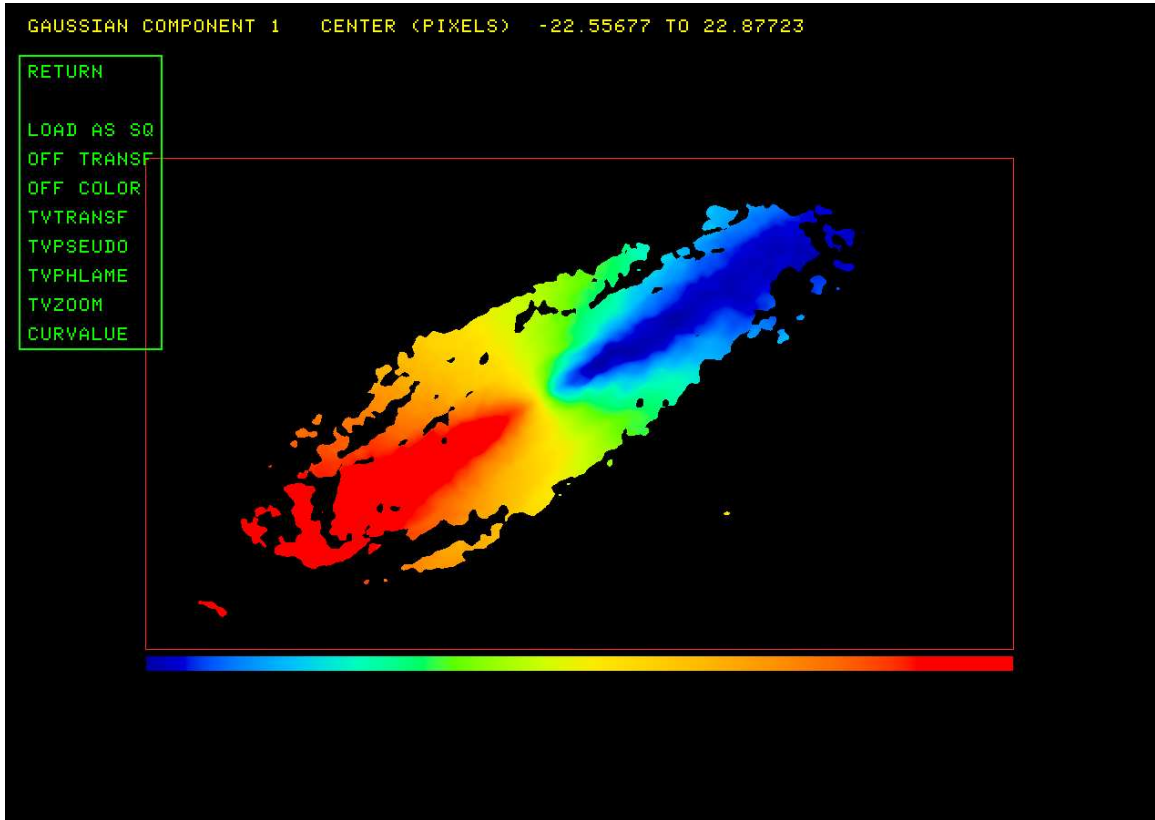
Figure 7: HI galaxy: center channel image interaction

RETURN	Return to the above menus, image stays displayed
LOAD AS SQ	Re-load image with square root transfer function
LOAD AS LG	Re-load image with log transfer function
LOAD AS L2	Re-load image with extreme log transfer function
LOAD AS LN	Re-load image with linear transfer function
OFF TRANSF	Turn off enhancement done with TVTRANSF
OFF COLOR	Turn off color enhancements
TVTRANSF	Black & white image enhancement
TVPSEUDO	Color enhancement of numerous sorts
TVPHLAME	Color enhancement of flame type, multiple colors
TVZOOM	Interactive zooming and centering of image
CURVALUE	Display value under cursor, mark pixels for list
SWAP 1-2	Swap solutions for components 1 and 2 interactively
NEXT WINDOW	Move to next window into large images

Two operations in this menu are different from the usual. CURVALUE provides the capability of selecting positions for the edit "list." During the CURVALUE operation position the cursor over the desired pixel and press buttons A, B, or C to add that pixel to the list. The SWAP $n-m$ operation uses a TV blotch operation like that in the $\mathcal{AIPS}$ verb TVSTAT and task BLSUM. You are to mark with a "blotch" region those pixels in the present image which are to have their solutions swapped with those of the selected component. Instructions will appear in the message window as you proceed. Begin by positioning the cursor at a pixel to be the first vertex of a connected sequence of vertices and press TV button A. Move to the next vertex and press button A again and repeat until you have marked all vertices for this region. Then press button D is you are done with this region or button C is you need to re-position one of the vertices. In this case, move the cursor to
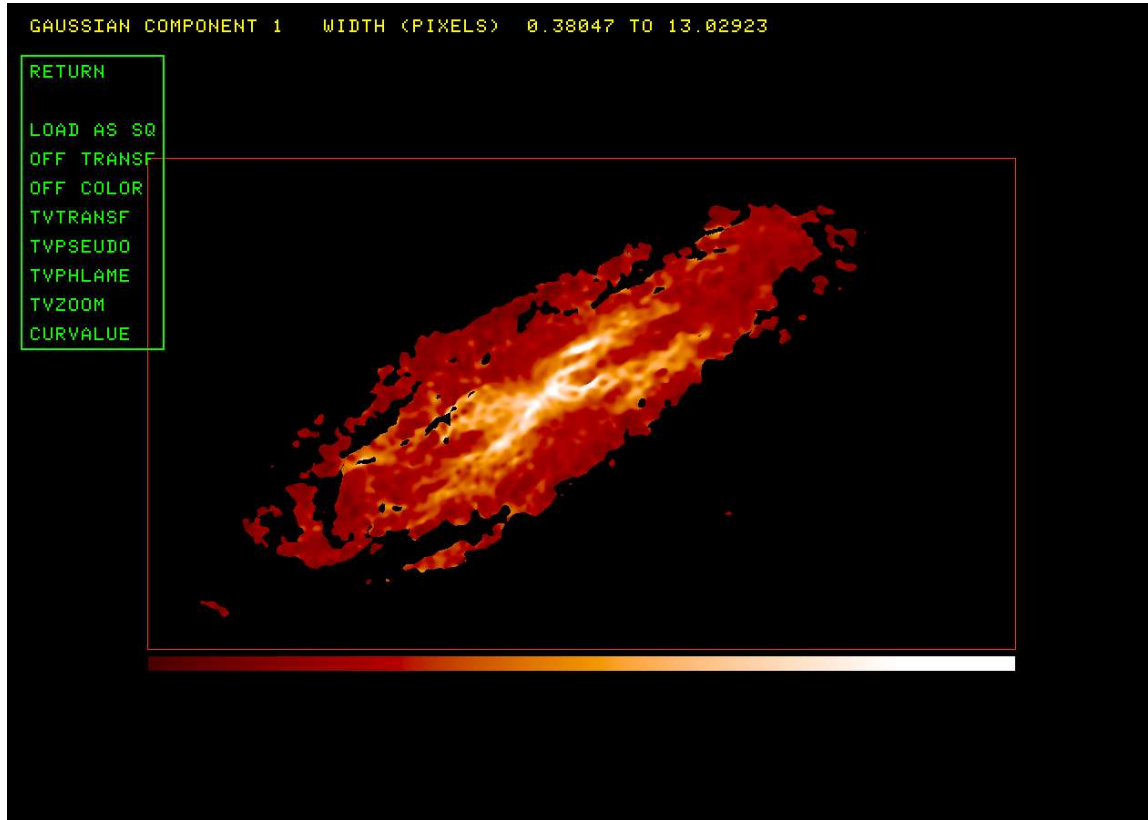
Figure 8: HI galaxy: width image interaction

the vertex to be re-positioned, press button `A` and drag the vertex to the corrected position. Press button `A` or `B` to fix that vertex and go on to reset another vertex or `D` to end this region and swap the solutions. You may do this as many times as needed.

When you have finished getting the images just the way you want them, you may write them out as $\mathcal{AIPS}$ image files. Select the `SET DOOUTPUT` option until its value, shown at the top of the screen, is what you want. In `XGAUS`, values 1 and 3 cause a residual image cube to be written, while values 2 and 3 cause images of the parameter values and their uncertainties to be written. The baseline and slope images and their uncertainties are given class `CONST`, `SLOPE`, `DCONST`, and `DSLOPE`, while the parameter images and their uncertainties get class `AMPL`n, `CENTR`n, `WIDTH`n, `DAMPL`n, `DCENT`n, and `DWIDT`n, and the flux and its uncertainty get class `FLUX`n, and `DFLUX`n, where n is the component number.

3   Zeeman splitting: `ZEMAN`

When an intrinsically unpolarized spectral line is emitted in the presence of a magnetic field, the right and left circular polarizations have their frequencies shifted in opposite directions by an amount proportional to the magnetic field (at least for modest magnetic fields). The traditional analysis of data to measure this splitting works only for very modest fields, those om which the separation of polarizations is a small fraction on the line width. In that case, the function that has been traditional is

$$V(x,y,z) = A(y,z)I(x,y,z) + 0.5B(y,z)\,\frac{dI(x,y,z)}{dx} \tag{2}$$

where $V(x, y, z)$ is the V Stokes polarization component. $I(x, y, z)$ is the unpolarized I Stokes component, x is the spectral axis value, (y, z) is the celestial coordinate value, and $A(y, z)$ and $B(y, z)$ are the parameters to be found by a linear least-squares method. The $\mathcal{AIPS}$ task ZEMAN offers this model, with two choices for the method by which the derivative of I with x is computed. If the total intensity spectrum has had XGAUS applied, another function may be used instead:

$$V(x, y, z) = A(y, z)I(x, y, z) + 0.5 \sum_{1}^{\text{NGAUSS}} B_i(y, z) \frac{dG_i(x, y, z)}{dx} \tag{3}$$

where $G_i(x, y, z)$ is the spectrum of the i'th Gaussian component and one solves for $A(y, z)$ and NGAUSS component frequency separations. This has been used to find the magnetic fields of components which have not been resolved spatially in the present data cube. Note that ZEMAN returns B in "channels;" any association with magnetic field values is left to the user,

3.1 Inputs

The inputs for ZEMAN are very similar to those for XGAUS. INNAME *et al.* specify the V polarization cube which must be in transposed form with frequency as the first axis. IN2NAME *et al.* specify the corresponding I polarization cube, similarly transposed. BLC and TRC define the spectral and celestial coordinate regions for the fit and FLUX gives the lower limit in the I image for the average of three consecutive channels for the spectra to be fit. INVERS specifies the input version of the ZE table in which the results are stored. Zero means to make a new one, other values mean to re-visit an existing solution. Adverb OPTYPE specifies which of the above formulæ are solved. OPTYPE = 'GAUS' syas to use the XG file IN2VERS attached to the input I polarization cube to solve using Equation 3. OPTYPE = '2SID' says to solve Equation 2, evaluating the derivative by

$$\frac{dI(x, y, z)}{dx} = 0.5\,[\,I(x+1, y, z) - I(x-1, y, z)\,]$$

while any other OPTYPE value says to evaluate the derivative with

$$\frac{dI(x, y, z)}{dx} = I(x, y, z) - I(x-1, y, z)$$

while solving Equation 2. DOOUTPUT controls what files are written — this may be changed interactively so leave it zero at this point. Set DOTV = 2 to use TV menus to prompt you. Even though the fitting operation is linear and so needs little or no guidance, it is best to watch what is happening so you should never set this adverb false. If all seems well, you can turn off the TV after watching a few of the solutions. DOMAX controls whether $A(y, z)$ is fit or taken to be zero. You may write out the input V cube with the fit $A(y, z$ removed, so DOMAX = -1 is appropriate when ZEMAN is applied to such a corrected cube. Of course, you could fit it again anyway, the result should be very close to zero. RMSLIMIT is an upper limit for the rms of a fit before the fit is viewed as "failed" which causes the TV and interaction to be turned back on after you have turned it off. You should get a good idea of an appropriate value from your initial uses of ZEMAN or from your knowledge of the noise in your data cube.

3.2 Fitting

3.3 Editing and output

4 Zeeman splitting: ZGAUS

ZEMAN is limited to the small-field approximation for Zeeman splitting in which the two polarizations are separated by only a small fraction of the line width. It is well-known that OH masers in the main lines at 1665 and 1667 MHz can be split by rather more than their full line width and intermediate cases can also

exist. For that reason, a new task tentatively called `ZGAUS` is needed. This task will simultaneously solve for the Gaussian components peak, center, and width and for a polarization separation term $s_i(y,z)$ and an I polarization leakage term $g(y,z)$. The correct equation to solve is

$$
I(x,y,z) = A + Bx + 0.5\sum_{1}^{\text{NGAUSS}} T_i e^{-\alpha(x-x_i-s_i/2)^2/\sigma_i^2}
$$

$$
+0.5\sum_{1}^{\text{NGAUSS}} T_i e^{-\alpha(x-x_i+s_i/2)^2/\sigma_i^2} \tag{4}
$$

$$
V(x,y,z) = g(y,z)I(x,y,z) + 0.5\sum_{1}^{\text{NGAUSS}} T_i e^{-\alpha(x-x_i-s_i/2)^2/\sigma_i^2}
$$

$$
-0.5\sum_{1}^{\text{NGAUSS}} T_i e^{-\alpha(x-x_i+s_i/2)^2/\sigma_i^2} \tag{5}
$$

Watch this space for future developments!

5 Polarization fitting

5.1 Inputs

5.2 Fitting

5.3 Editing and output

6 Model creation

6.1 Visibility data models

6.2 Image data models